Published in IET Signal Processing
Received on 6th November 2008
Revised on 24th June 2009
doi: 10.1049/iet-spr.2008.0224

ISSN 1751-9675

Adaptive grid risk-sensitive filter for non-linear problems

S. Bhaumik[1] *M. Srinivasan*[2] *S. Sadhu*[3] *T.K. Ghoshal*[3]

[1]*GE India Technology Centre, JFWTC, Bangalore 560 066, India*
[2]*Koneru Laksmaiha College of Engineering, Vadeswaram, Guntur District, AP, India*
[3]*Department of Electrical Engineering, Jadavpur University, Kolkata 700 032, India*
E-mail: smita_sadhu@rediffmail.com

Abstract: A novel adaptive grid-based method has been proposed for risk-sensitive state estimation in non-linear non-Gaussian problems. The algorithm, which is based on point-mass approximation, is called the adaptive grid risk-sensitive filter (AGRSF). Although risk-sensitive estimators have been known to be robust compared to their risk-neutral counterparts, the implementation of risk-sensitive filters (RSFs) is almost impossible except for very trivial systems like linear Gaussian systems. The existing extended risk-sensitive filter (ERSF) fails to take care of non-Gaussian problems or severe non-linearities. Recently, other variants of RSFs have been proposed for extending the range of applications of risk-sensitive techniques. The AGRSF has been formulated to act as a benchmark and aid in the validation of other RSFs. The algorithm uses a modified form of information state-based recursive relation and provides guidelines for the adaptive choice of grid points to improve the numerical efficiency. The developed filter has been applied to a single-dimensional non-linear poorly observable system and a non-linear two-dimensional bearing only tracking problem. The convergence of the algorithm has been shown by simulation. The estimation efficiency and computational load of AGRSF has been compared with other RSFs.

1 Introduction

Although the robustness of risk-sensitive filters (RSFs) has been known for over a decade, interest in it has been renewed only recently [1]. This has been primarily because of the difficulty in implementing these filters for non-linear systems, as these filters admit closed-form solutions only for a very limited class of systems. The properties of RSFs have been summarised in [2]. A more general framework compared to the earlier publications [2, 3] has been used in [1]. The concept of risk-sensitive estimation (RSE) is equally applicable to linear as well as non-linear problems. The closed-form solution exists only for linear Gaussian case. For non-linear case an extended Kalman filter (EKF)-like framework for RSE, called the extended risk-sensitive filter (ERSF) has been given in [4]. However, all the limitations of EKF, including smoothness requirement for the functions and noise Gaussianity restriction, are inherited by the ERSF. Until some recent work, where some newer variants like central difference risk-sensitive filter (CDRSF) [5], risk-sensitive unscented Kalman filter (RSUKF) [6] and risk-sensitive particle filter (RSPF) have been proposed [7, 8], there had been no general method for solving RSF problems for non-Gaussian, non-linear cases.

The adaptive grid risk-sensitive filter (AGRSF), based on point-mass approximation [9], is a novel method being proposed in this paper for non-linear non-Gaussian risk-sensitive state estimation problems. In the process of proposing the AGRSF, the recursive solution of general RSE has been stated. This particular form is amenable to a more efficient AGRSF implementation compared to the previously available forms.

.The development of the adaptive grid-based approach was primarily motivated by the need to validate other risk-sensitive filters like CDRSF, RSUKF and RSPF. Although a grid-based implementation is known to be computationally inefficient, the straightforward mapping between the general formulation given in [1] and the algorithmic implementation of grid-based filter allows an enhanced confidence in the implemented algorithm. Grid adaptation methods enhance the numerical efficiency of the filter. Thus the AGRSF can act as a benchmark and aid in the validation of other RSFs. Further, as in the RSPF, the AGRSF provides easier visualisation of the prior and posterior densities that may be effectively used to diagnose implementation problems and shortcomings of alternative risk-sensitive filtering methods.

In this paper, which is an expanded version of [10, 11], convergence of the proposed filter has been tested by simulation with increasing number of support grid points. The filter has been applied to a one-dimensional non-linear poorly observable system and a standard non-linear two-dimensional bearing only tracking problem. The performance of the AGRSF is compared with the ERSF. The computational load of proposed filter has been compared with RSPF and the results have been tabulated.

It may be mentioned that the application of AGRSF requires selecting the risk-sensitive parameter μ. The issues involved in selecting μ have been discussed adequately in

[2, 3]. However, the algorithm for AGRSF is not dependent on selection of μ.

2 RSE problem

Consider a general (non-linear) signal model consisting of equations for the state $\boldsymbol{x}_k \in R^n$ and measurement $\boldsymbol{y}_k \in R^p$ with additive, uncorrelated noise $\boldsymbol{w}_k \in R^n$, $\boldsymbol{v}_k \in R^p$ of known statistics at the instance $k = \{0, 1, 2, 3, \ldots, n\}$

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k) + \boldsymbol{w}_k \tag{1}$$

$$\boldsymbol{y}_k = \boldsymbol{g}(\boldsymbol{x}_k) + \boldsymbol{v}_k \tag{2}$$

The vectors $\boldsymbol{f}(\boldsymbol{x}_k)$ and $\boldsymbol{g}(\boldsymbol{x}_k)$ are general (without any assumption of smoothness) non-linear functions of $\boldsymbol{x}_k$ and k. The initial state $\boldsymbol{x}_0$ is uncorrelated with the noises mentioned above, but has a known probability density distribution p_0.

From expression 4.2 in [1], when a quadratic function is substituted for the more general convex function $\rho(.)$, a two-parameter risk-sensitive squared error cost function at the time instance k may be defined as

$$J_{\mathrm{RS}}(\boldsymbol{\varsigma}, k) = E\left[\exp\left(\mu_1 \sum_{i=0}^{k-1} (\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i)^{\mathrm{T}}(\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i) + \mu_2 (\boldsymbol{x}_k - \boldsymbol{\varsigma})^{\mathrm{T}}(\boldsymbol{x}_k - \boldsymbol{\varsigma})\right)\right] \tag{3}$$

where $\hat{\boldsymbol{x}}_i$'s are the optimum estimated values of state variable for past steps $i \in \{0, 1, 2, 3 \ldots k-1\}$. The constant parameters μ_1 and μ_2 are called risk-sensitive parameters. The current optimum estimate $\hat{\boldsymbol{x}}_k$ is obtained by finding the optimum value of $\boldsymbol{\varsigma}$, which minimises $J_{\mathrm{RS}}(\boldsymbol{\varsigma}, k)$ that is

$$\hat{\boldsymbol{x}}_k = \arg\min_{\boldsymbol{\varsigma}} J_{\mathrm{RS}}(\boldsymbol{\varsigma}, k) \tag{4}$$

3 Recursive solution of the RSE problem

The recursive solution of the RSE problem may be realised, generally, as a two-step process:

1. formulation of a recursive relation of an information state [1, 10, 11], which is updated in each time step, and
2. optimisation of a cost function involving the information state.

It can be easily shown [8] that the solution of the RSE problem may be obtained from the following recursive relations

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \arg\min_{\boldsymbol{\varsigma}} \int_{-\infty}^{+\infty} \exp[\mu_2(\boldsymbol{x}_{k+1} - \boldsymbol{\varsigma})^{\mathrm{T}} \times (\boldsymbol{x}_{k+1} - \boldsymbol{\varsigma})]\alpha_{k+1}\, \mathrm{d}\boldsymbol{x}_{k+1} \tag{5}$$

where

$$\alpha_{k+1} = p(\boldsymbol{y}_{k+1}|\boldsymbol{x}_{k+1}) \int_{-\infty}^{+\infty} \exp[\mu_1(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)] \times \alpha_k p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)\, \mathrm{d}\boldsymbol{x}_k \tag{6}$$

The above estimation equation is specifically applicable for posterior estimation where the estimation is performed after receiving the measurement. For prior estimates, relations similar to that given in [1] should be used.

4 Adaptive grid formulation

The integrations given in relations (5) and (6), may not, in general be performed in closed form for non-linear or non-Gaussian case. A convenient numerical technique to carry out this integration is point-mass approximation, which is used in grid-based filters [9, 12].

4.1 Representing probability densities as point mass on grid points

The problem is to approximate at time step k, a probability distribution $p(\boldsymbol{x})$ defined everywhere in state space denoted by $\boldsymbol{x}$. Let it be known a priori that outside the region Γ bounded by $\boldsymbol{x} = \boldsymbol{X}_0$ and $\boldsymbol{X}_{\mathrm{f}}$, the probability density is negligibly small.

The region of interest Γ in the state space is subdivided into N grid points with more grid points in the region where the solution is expected. At time step k, let the non-trivial grid points, also called support points, in state space be denoted as $\boldsymbol{x}_k^j$ for $j = 1, 2, \ldots, N$. The idea of grid and support points is shown in Fig. 1.

In the case of a single state variable, from prior knowledge, one can set the minimum and maximum values and divide the interval into N grids. The central point between grid line 0 and grid line 1 may be called as the grid point 1. Similarly grid point j will be in the centre of grid lines $j-1$ and j.

In a two-dimensional case, state variables may be divided into N_1 and N_2 grid lines, creating $N = N_1 N_2$ grid points. The area bounded by the adjacent grid lines will be the corresponding grid hyper surface. The grid points need to be numbered systematically.

For grid point $\boldsymbol{x}_k^j$, the probability mass is m_k^j and $\sum_{j=1}^{N} m_k^j = 1$. The probability density may be approximated by point mass as $p(\boldsymbol{x}_k) \simeq \sum_{i=1}^{N} m_k^i \delta(\boldsymbol{x}_k - \boldsymbol{x}_k^i)$.

Although the information state α_k is not a probability density, it does represent a density. It may therefore be approximated likewise. Support points for α_k are generally chosen to be the same as for $p(\boldsymbol{x}_k)$. It may be seen easily that estimation is unaffected if α_k is multiplied by a constant factor. It is often advantageous to normalise the distribution of α_k.

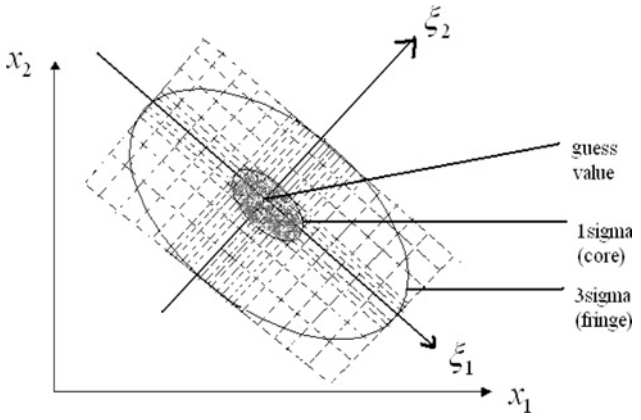

Fig. 1 *Idea of grid and support points*

doi: 10.1049/iet-spr.2008.0224

4.2 Grid-based RSE

The adaptive grid algorithm has two distinct aspects, namely, the selection of grid points (and grid axes) by adaptation discussed in the next section and the estimation algorithm using the chosen set of grid points (also called supports) using point-mass approximation. Unlike conventional point-mass filters [9, 12] where the probability distribution of the states is discretised, in the AGRSF, the analogous quantity, that is, the information state is discretised. This approach is similar to the 'probabilistic interpretation' used in [8].

For the set of chosen grid points, the initial state and the information state are discretised into the point-mass form as

$$p_{0|0}(\boldsymbol{x}_0) \simeq \sum_{i=1}^{N} m_0^i \delta(\boldsymbol{x}_0 - \boldsymbol{x}_0^i) \quad \text{and}$$

$$\alpha_0(\boldsymbol{\varsigma}) \simeq \sum_{i=1}^{N} \alpha_0^i \delta(\boldsymbol{\varsigma} - \boldsymbol{x}_0^i)$$

For the given set of grid points, the information state is updated from the kth to the $(k+1)$th step as

$$\alpha_{k+1}^i(\boldsymbol{x}_{k+1}^i) = p_v(\boldsymbol{y}_{k+1} - \boldsymbol{g}(\boldsymbol{x}_{k+1}^i)) \sum_{j=1}^{N} p_w(\boldsymbol{x}_{k+1}^i - \boldsymbol{f}(\boldsymbol{x}_k^j))\alpha_k^j(\boldsymbol{x}_k^j) \times \exp\{\mu_1(\boldsymbol{x}_k^i - \hat{\boldsymbol{x}}_k)^{\mathrm{T}}(\boldsymbol{x}_k^i - \hat{\boldsymbol{x}}_k)\} \tag{7}$$

The above expression is the discretised form of (6) using the theory of approximate grid-based filter [12]. The alpha masses are normalised so that the sum is unity. The objective function to be minimised is then approximated likewise, for any arbitrary point $\boldsymbol{\varsigma}$ as

$$J_{\mathrm{RS}}(\boldsymbol{\varsigma}, k) \simeq \sum_{j=1}^{N} \exp[\mu_2(\boldsymbol{x}_k^j - \boldsymbol{\varsigma})^{\mathrm{T}}(\boldsymbol{x}_k^j - \boldsymbol{\varsigma})]\alpha_k^j$$

The risk-sensitive state estimate is obtained by determining the optimising solution for the objective function. Out of several possible methods for determining the optimum, the solution for the zero value of the partial derivative of the objective function has been used in the present work

$$\frac{\partial}{\partial \boldsymbol{\varsigma}} J_{\mathrm{RS}}(\boldsymbol{\varsigma}, k) = 0 \Rightarrow \sum_{j=1}^{N} (\boldsymbol{x}_k^j - \boldsymbol{\varsigma})\alpha_k^j \exp[\mu_2(\boldsymbol{x}_k^j - \boldsymbol{\varsigma})^{\mathrm{T}}(\boldsymbol{x}_k^j - \boldsymbol{\varsigma})]\alpha_k^j = 0$$

If the distribution is Gaussian, the solution of this equation is the mean value.

4.3 Adaptation of grids

In conventional grid-based filters, use of fixed support points leads to lack of computational efficiency. Either many such support points are redundant, contributing only to the numerical load without improving the accuracy of the estimates, or else, if too coarse a grid is chosen, estimation accuracy suffers. In adaptive grid-based method, the number of wasted grid points is minimised by adaptive selection of grid points. The following pragmatics has been proposed in the present work.

- The grid points would have to be constrained within a hyper volume of the (information) state space, called the region of interest. The region, bounded by the minimum and maximum values (also called the extent or span) in each axis should be made as small as possible.
- The span should be divided into a number of intervals consistent with the required resolution.
- A non-uniform density grid should be chosen, so that the grid density is high near the expected estimate and low at the extremities.
- The direction of grid axes should be properly aligned to reduce wasted points. This is exemplified in Fig. 1 where the grid axes are inclined with respect to the state variable axes to reduce wasted support.

4.3.1 Obtaining approximate probability distribution: Application of the pragmatics given in the previous section requires the use of an approximate probability distribution. In the proposed algorithm, EKF has been used to obtain such an approximate distribution. The EKF provides only an inaccurate estimate, which, however, has been found to be adequate for grid adaptation to enhance the efficiency of AGRSF.

From the approximate mean and covariance, provided by the above approximation, we obtain not only the boundaries of the grid but also good orientation of the grid axes as explained in the following sections.

4.3.2 Obtaining approximate boundaries of the region of interest: The EKF-based method provides the mean and covariance of the concerned distribution with a Gaussian assumption. The mean provides a pragmatic location for the centre of the grid axis system, the eigenvectors of the covariance matrix provide the direction of the axes and from the corresponding eigenvalues, one can decide on the extent of the grids (boundaries of the region of interest), as explained below.

We first define the m-sigma contour (m-sigma hyper surface for higher dimension) where m is any real number, for an n-dimensional distribution with a covariance matrix $\boldsymbol{P}$, as the locus of the points satisfying the n-dimensional quadratic equation $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{x} = m^2$, in state space.

To appreciate the approach, let us consider a 2-D problem where the covariance matrix is 2×2 and

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}$$

The locus of the m-sigma contour is then given by the equation

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathrm{T}} = m^2 \Rightarrow \beta_{11}(x_1)^2 + (\beta_{21} + \beta_{12})x_1x_2 + \beta_{22}(x_2)^2 = m^2$$

This is clearly the equation of an ellipse. If the distribution is Gaussian-like with normal tails, a three-sigma contour will contain almost all the points with significant density. (For a heavier tailed distribution, a higher value of m may be necessary.) The task is then converted to finding the extremities of the three-sigma contour and putting a rectangular grid to fit the contour.

The intercepts of the sigma contour at the x_1 and x_2 axes would be $\pm m\sqrt{1/\beta_{11}}$, $\pm m\sqrt{1/\beta_{22}}$, respectively. If the axes of the ellipse coincide with the state variables, the

intercepts are the semi-major/minor axes and the dimension of the enclosing box is easily obtained. The knowledge of the intercept points is useful to define the extremities of the region of interest if the axes of the ellipse coincide with the state variables.

4.3.3 Aligning the axes: For a more general case the extremities of the region of interest may be obtained if the axes are chosen along the major and minor axes of the m-sigma contour and the intercepts on these axes are used as the extremities (as shown in Fig. 1).

It may be easily shown that the principal axes of the ellipse/ellipsoid $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{-1}\boldsymbol{x} = m^2$ would be along the eigenvectors of the symmetric matrix $\boldsymbol{P}^{-1}$. The extremities of the ellipse would remain at $\pm m\sqrt{1/\beta_{11}}$, $\pm m\sqrt{1/\beta_{22}}$.

4.3.4 Selecting grid intervals: The next level of adaptation ensures that there would be enough support points in the core area, where a solution is expected, and adequate points to represent the density in the fringe/tail areas. From a priori knowledge about the gross nature of the posterior distribution (like whether it is of heavy tailed type, leptokurtic and so on) one can choose the core region and the fringe region appropriately. Typically, the core area may be taken as the one-sigma boundary around the estimated state. The fringe area could be the three-sigma boundary, leaving out the core. The idea is to keep approximately 60% of the support points in the core area and the rest for the fringe. The basic heuristics are further improved by taking a smoothly varying grid region as one goes from the core area to the fringe.

5 AGRSF algorithm

In this algorithm, posterior estimation is assumed.

Step 1: Grid template generation

- Select the number of grid intervals in each axis, N^i, that is adequate for typical resolution requirements, where i is the grid index. Total number of grid points $N \leq (N^i)^n$, where n is the dimension of state space.
- Choose the outer or fringe area boundary/extent descriptor m. Choice of descriptor m may be obtained after preliminary experimentation starting with a value of $m = 4$.
- Divide the total extent $2m$ into N^i intervals. The intervals will be symmetrical about the centre and spaced according to the non-uniform spacing strategy described earlier.
- Generate the support points systematically in the hyper space. Let z^i be a typical support point, with coordinates $(z_1^i, z_2^i, z_3^i, \ldots, z_j^i, \ldots, z_n^i)$. The support point z^i would be located at the centroid of the grid cell formed by the grid intervals of each dimension, such that each of the coordinates z_j^i is at the midpoints of corresponding grid intervals.
- Generate the corresponding grid cell volumes $\tilde{q}^i$ for each support point and normalise the $\tilde{q}^i$'s to produce q^i.
- Store the $N \times n$ coordinate points in the Z-plane as the grid template.

Step 2: Filter initialisation

- Set $k = 0$.
- Set initial conditions $\hat{\boldsymbol{x}}_{0|0}$, and $\boldsymbol{P}_0$ which are assumed or given.
- Obtain the grid alignment transform $\boldsymbol{S}_0$ from the initial covariance $\boldsymbol{P}_0$, using $\boldsymbol{S}_0\boldsymbol{S}_0^{\mathrm{T}} = \boldsymbol{P}_0$.
- Obtain the initial support points $\boldsymbol{x}_0^i$, $i = 1, 2, 3, \ldots, N$, from $\boldsymbol{x}_0^i = \boldsymbol{S}_0\boldsymbol{z}^i + \hat{\boldsymbol{x}}_{0|0}$.
- Compute the weights $\tilde{w}_0^i$, corresponding to $\boldsymbol{x}_0^i$ from $\boldsymbol{P}_0$, by $\tilde{w}_0^i = q^i\rho^i$, where q^i is the grid cell volume and ρ^i is the average local density in the cell, obtainable from the known probability distribution $p(\boldsymbol{x}_{0|0})$. The weights are then normalised and the normalised weights are denoted as w_0^i.

Step 3: Computation of risk-sensitive estimate for $(k+1)$th step.

- Obtain an approximate value of the posterior mean $\boldsymbol{\xi}_{k+1} \simeq \hat{\boldsymbol{x}}_{k+1|k+1}$ at $(k+1)$th step by EKF. Such filters use the mean $\hat{\boldsymbol{x}}_k$ and covariance $\boldsymbol{P}_k$ as starting point and generate the approximate posterior error covariance $\tilde{\boldsymbol{P}}_{k+1}$ as a by-product.
- Obtain the lower triangular matrix square root for the approximate posterior covariance $\boldsymbol{S}_{k+1}$ from $\tilde{\boldsymbol{P}}_{k+1}$, using $\boldsymbol{S}_{k+1}\boldsymbol{S}_{k+1}^{\mathrm{T}} = \tilde{\boldsymbol{P}}_{k+1}$.
- Obtain the new support points $\boldsymbol{x}_{k+1}^i$, $i = 1, 2, 3, \ldots, N$ from the grid template, $\boldsymbol{S}_{k+1}$ and approximate posterior mean $\boldsymbol{\xi}_{k+1}$, by using the relation $\boldsymbol{x}_{k+1}^i = \boldsymbol{S}_{k+1}\boldsymbol{z}^i + \boldsymbol{\xi}_{k+1}$.
- Obtain the weights $\tilde{w}_{k+1|k}^i$ corresponding to $\boldsymbol{x}_{k+1}^i$ from $\tilde{w}_{k+1|k}^i = q^i \sum_{j=1}^N \boldsymbol{w}_k^i(\boldsymbol{x}_k^j) \exp[\mu_1(\boldsymbol{x}_k^i - \hat{\boldsymbol{x}}_k)^{\mathrm{T}}(\boldsymbol{x}_k^i - \hat{\boldsymbol{x}}_k)] p_w(\boldsymbol{x}_{k+1}^i - \boldsymbol{f}(\boldsymbol{x}_k^j))$. Obtain the weights $\tilde{w}_{k+1|k+1}^i$ corresponding to $\boldsymbol{x}_{k+1}^i$ from the relation $\tilde{w}_{k+1|k+1}^i = \tilde{w}_{k+1|k}^i \times p_v(\boldsymbol{y}_{k+1} - \boldsymbol{g}(\boldsymbol{x}_{k+1}^i))$. Normalise the set of weights $\tilde{w}_{k+1|k+1}^i$ and obtain the normalised weights w_{k+1}^i.
- Obtain the expected value of the posterior estimate (optimum in minimum mean square sense) from $\hat{\boldsymbol{x}}_{k+1|k+1} = (1/N)\sum_{i=1}^N w_{k+1}^i \boldsymbol{x}_{k+1}^i$.
- Compute the updated posterior covariance matrix $\boldsymbol{P}_{k+1}$ by using the relation $\boldsymbol{P}_{k+1} = (1/N)\sum_{i=1}^N w_{k+1}^i (\boldsymbol{x}_{k+1}^i - \hat{\boldsymbol{x}}_{k+1|k+1})(\boldsymbol{x}\ k+1^i - \hat{\boldsymbol{x}}_{k+1|k+1})^{\mathrm{T}}$.

Step 4: Recursion

- Compute estimates for the subsequent steps by repeating step 3 as above for the required number of time steps.

Note 1: In this algorithm, the a posteriori value of the state has been estimated. Similar algorithm can be formulated for obtaining the a priori risk-sensitive estimate.
Note 2: In this work, the probability density function obtained from EKF has been used as a proposal density function to draw the support points. Other filtering algorithms (like UKF, CDF, ERSF, CDRSF and so on) can be used to generate the proposal for AGRSF.

6 Case studies

The AGRSF has been cross-validated against the closed-form solution [that is risk-sensitive Kalman filter (RSKF)] for a one-dimensional linear Gaussian problem and the results of validation have been presented elsewhere [10]. In the current paper, we consider two non-linear problems. In the first case study, the AGRSF is applied to a one-dimensional non-linear poorly observable system [8, 13]. The second case study provides results for the two-dimensional bearing only tracking problem.

6.1 Case study-I

6.1.1 Problem statement: This problem first appeared in [13] and has been revisited with a different measurement equation in [8]. This one-dimensional non-linear signal

doi: 10.1049/iet-spr.2008.0224

process is given by the state equation

$$x_k = x_{k-1} + \Delta t\, 5x_{k-1}(1 - x_{k-1}^2) + w_k, \quad w_k \sim \mathcal{N}(0, b^2\Delta t) \tag{8}$$

and the measurement equation

$$y_k = \Delta t x_k(1 - 0.5x_k) + v_k, \quad v_k \sim \mathcal{N}(0, d^2\Delta t) \tag{9}$$

where $\sim \mathcal{N}(\eta, \theta)$ denotes a sampling process from normal distribution with mean $= \eta$ and covariance $= \theta$; the sampling time is $\Delta t = 0.01$ sec, the constants b and d are $b = 0.5$ and $d = 0.5$. Given also that the true initial value is $\boldsymbol{x}_0 = \boldsymbol{x}(0) = -0.2$, filter initial state $\hat{\boldsymbol{x}}_{0|0} = 0.8$, filter initial error covariance is $\boldsymbol{P}_{0|0} = 2$ and value of risk-sensitive parameter (μ_1) is 0.0756.

6.1.2 Comparison with ERSF: The problem has been solved using AGRSF with 100 grid points and the result is compared with that obtained from ERSF. The truth and estimated value of state obtained from the above-mentioned filters is shown in Fig. 2*a* for a representative run. It is seen that for this particular run the ERSF fails to track the truth whereas the AGRSF tracks the truth well.

The RMS errors (RMSE) of AGRSF and ERSF have been compared using 1000 Monte Carlo runs shown in Fig. 2*b*. It can be noted that the RMSE of ERSF is quite high compared to RMSE of AGRSF. It may be pointed out that in the non-linear problem used, even though the noises are Gaussian, the prior and posterior densities in the estimation steps may be non-Gaussian. A typical case is shown in Fig. 3, where the non-Gaussianity is obvious. In fact, the distribution is bi-modal in nature. Because of the ability of the AGRSF to deal with such non-Gaussian distributions, its performance is much better than that of ERSF.

6.1.3 Comparison with RSPF: In [11], which was a preliminary version of the current work, results had been provided to show the relative performances of the RSPF and AGRSF for this particular problem. Results in [11] show that although the AGRSF is noticeably slower than the RSPF, the RMSE of both the RSPF and AGRSF settle to (approximately) the same value. This helps in validating both the RSPF and AGRSF algorithms.

As a part of the present study, computational loads for AGRSF and RSPF have been studied and compared. For the comparison to be meaningful, one must use appropriate number of grids and particles that provide similar accuracy of estimation. It is observed that for this particular problem, 100 grid points and 1000 particles are sufficient for AGRSF and RSPF, respectively, to obtain equivalent results. The time for a single run of AGRSF (with different number of grid points) and the time for a single run of RSPF (for different number of particles) in Matlab software in a P-IV (2.66 GHz) processor are summarised in Table 1. From the table it is seen that the computational efficiency of RSPF is substantially better than the AGRSF. This, however, is not unexpected. The superior numerical efficiency of particle filter over (ordinary) point-mass methods is well known as the number of computations in grid-based filter is of the order of N-squared (N being the number of support points) whereas the same for particle-based filter is of the order of N [14]. The table confirms a similar trend in the risk-sensitive context.

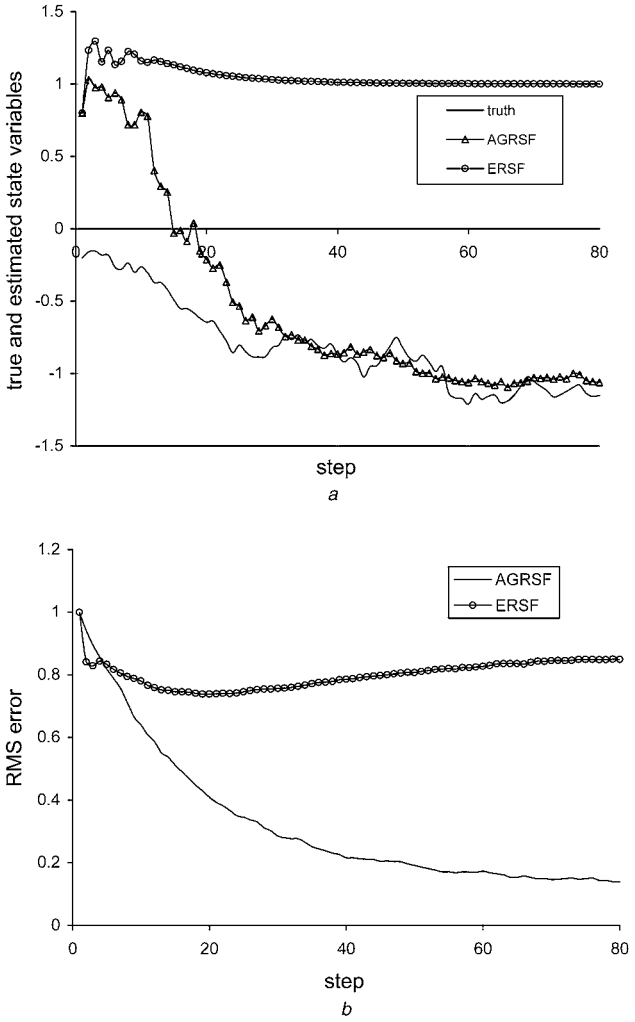

Fig. 2 *Filter performance for Case study-I*

a Truth and estimated values of state obtained from ERSF and AGRSF
b RMSE of ERSF and AGRSF for 1000 MC run

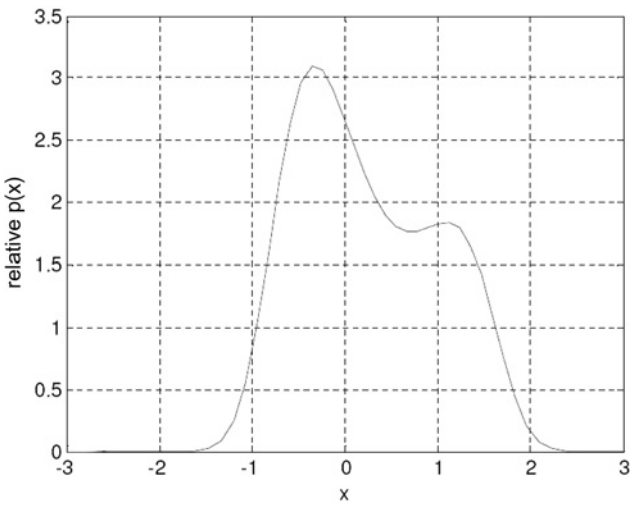

Fig. 3 *Posterior probability density for Case study-I at the 15th step*

6.2 Case study-II

This example studies a bearing only tracking problem with linear system and non-linear measurements. A detailed

Table 1 Run time for a single run of RSPF and AGRSF for different no. of grid points and particles

AGRSF		RSPF	
No. of grids	Run time(s)	No. of particles	Runtime(s)
100	80.375	1000	21.06
200	312.407	2000	56.703
400	1265.3	5000	252.844

description of the problem statement can be found in [15, 16]. This problem has been chosen as it is a practical problem where many conventional filters are known to lose track.

Here, the tracking platform moves approximately parallel to the target with approximately constant velocity, given by the following discrete time equations

$$\boldsymbol{x}_{\mathrm{p}}(k) = \bar{\boldsymbol{x}}_{\mathrm{p}}(k) + \Delta\boldsymbol{x}_{\mathrm{p}}(k), \quad k = 0, 1, \ldots, n_{\mathrm{step}} \tag{10}$$

$$\boldsymbol{y}_{\mathrm{p}}(k) = \bar{\boldsymbol{y}}_{\mathrm{p}}(k) + \Delta\boldsymbol{y}_{\mathrm{p}}(k), \quad k = 0, 1, \ldots, n_{\mathrm{step}} \tag{11}$$

where $\bar{\boldsymbol{x}}_{\mathrm{p}}(k)$ and $\bar{\boldsymbol{y}}_{\mathrm{p}}(k)$ are the (known) average platform position coordinates, $\Delta\boldsymbol{x}_{\mathrm{p}}(k) \sim \mathcal{N}(0, T^2)$ and $\Delta\boldsymbol{y}_{\mathrm{p}}(k) \sim \mathcal{N}(0, T^2)$ are the mutually independent white noise sequences where T (=1 s) is the sampling time.

The mean position coordinates of the platform are

$$\bar{\boldsymbol{x}}_{\mathrm{p}}(k) = 4kT \quad \text{and} \quad \bar{\boldsymbol{y}}_{\mathrm{p}}(k) = 20$$

The target motion in the $\boldsymbol{x}$ direction is given by

$$\begin{bmatrix} \boldsymbol{x}_1(k+1) \\ \boldsymbol{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1(k) \\ \boldsymbol{x}_2(k) \end{bmatrix} + \begin{bmatrix} T^2/2 \\ T \end{bmatrix} w(k) \tag{12}$$

where $\boldsymbol{x}_1(k)$ is the position along the x-axis in metres, $\boldsymbol{x}_2(k)$ is the velocity in m/s and $w(k) \sim \mathcal{N}(0, q)$. The (unknown) true initial condition is $\underline{\boldsymbol{x}}(0) = \begin{bmatrix} 80 \\ 1 \end{bmatrix}$, and the known noise variance is $q = 0.01\,\mathrm{m}^2/\mathrm{s}^4$.

The measurement equation (in bearing coordinate) is given as

$$z_{\mathrm{m}}(k) = \tan^{-1} \frac{\boldsymbol{y}_{\mathrm{p}}(k)}{\boldsymbol{x}_1(k) - \boldsymbol{x}_{\mathrm{p}}(k)} + \boldsymbol{v}_{\mathrm{s}}(k) \tag{13}$$

where $\boldsymbol{v}_{\mathrm{s}}(k) \sim \mathcal{N}(0, r_{\mathrm{s}})$ and $r_{\mathrm{s}} = (3^\circ)^2$.

The effect of platform motion noises has been approximated as an additive noise by using the method described in [15, 16].

The filter initial velocity estimate is selected as $\hat{\boldsymbol{x}}_2(0) = 0$. Also, $\boldsymbol{P}_{22}(0) = 1$, $\boldsymbol{P}_{12}(0) = \boldsymbol{P}_{21}(0) = 0$.

Fig. 4 shows the plot of the absolute error in position (metre that is m) for a typical run with ERSF and AGRSF. In Fig. 5, it is shown with the help of simulation that for $\mu = 0.000001$, the absolute value of the mean position error converges as the number of grid points is increased.

7 Discussion and conclusion

An adaptive grid-based approximation of the recursive relation for solution of the RSE problem has been presented. The numerical efficiency of grid-based filtering has been improved substantially by using a number of grid adaptation mechanisms. AGRSF does not need random sampling and re-sampling as in the particle filter version. An important issue in using AGRSF is selecting the number of support points. This has to be obtained experimentally by trading off accuracy against computation time. The proposed non-linear risk-sensitive filter has been applied to a non-linear single dimensional poorly observable system and then a practical bearing only tracking problem. The convergence of the adaptive grid filter has been shown using simulation. Results of comparison show that the RMS error settling performance of the AGRSF is much better than what is obtainable from the ERSF and the performance of the former is almost similar to that of the RSPF. Although the RMS error settling rate of the AGRSF is marginally slower compared to the RSPF, both these filters settle to around the same final value. With increased number of grids, the accuracy of the AGRSF can be improved even further, though this would deteriorate its numerical efficiency. Thus the AGRSF would be suitable as a benchmark for off-line performance validation of other non-linear filters.

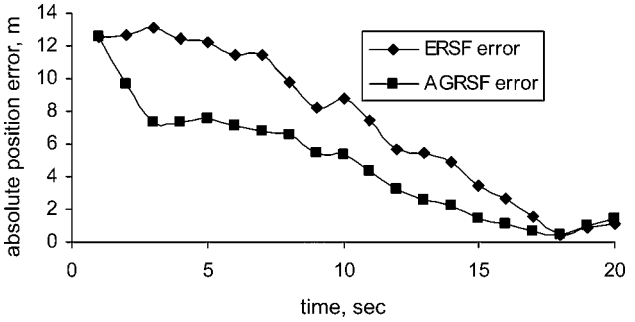

Fig. 4 *Absolute error in position for typical run in BOT problem*

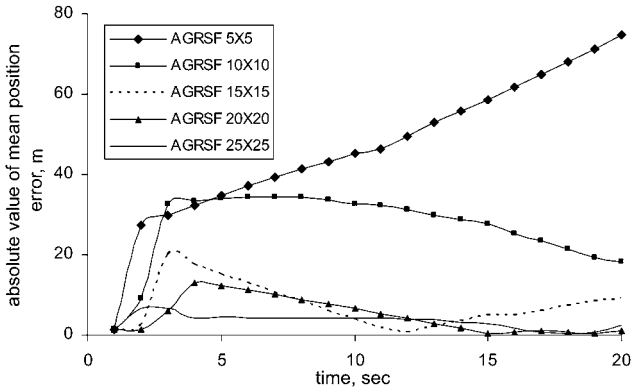

Fig. 5 *Convergence of AGRSF with number of grid points* (In figure, the value after AGRSF indicates the dimensions of grid used in each case)

8 References

1 Boel, R.K., James, M.R., Petersen, I.R.: 'Robustness and risk-sensitive filtering', *IEEE Trans. Autom. Control*, 2002, **47**, (3), pp. 451–460
2 Banavar, R.N., Speyer, J.L.: 'Properties of risk-sensitive filters/estimators', *IEE Proc. Control Theory Appl.*, 1998, **145**, (1), pp. 106–112
3 Moore, J.B., Elliott, R.J., Dey, S.: 'Risk-sensitive generalizations of minimum variance estimation and control', *J. Math. Syst. Estimation Control*, 1997, **7**, (1), pp. 1–15

doi: 10.1049/iet-spr.2008.0224

4 Jayakumar, M., Banavar, R.N.: 'Risk-sensitive filters for recursive estimation of motion from images', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1998, **20**, (6), pp. 659–666

5 Sadhu, S., Srinivasan, M., Bhaumik, S., Ghoshal, T.K.: 'Central difference formulation of risk-sensitive filter', *IEEE Signal Proc. Lett.*, 2007, **14**, (6), pp. 421–424

6 Bhaumik, S., Sadhu, S, Ghoshal, T.K.: 'Risk-sensitive formulation of unscented Kalman filter', *IET Control Theory Appl.*, 2009, **3**, (4), pp. 375–382

7 Sadhu, S., Doucet, A.: 'Particle methods for risk-sensitive filtering'. Proc. IEEE INDICON, December 2005, pp. 423–426

8 Sadhu, S., Bhaumik, S., Doucet, A., Ghoshal, T.K.: 'Particle method based formulation of risk-sensitive filter', *Signal Process.*, 2009, **89**, (3), pp. 314–319

9 Šimandl, M., Královec, J., Söderström, T.: 'Advanced point-mass method for nonlinear state estimation', *Automatica*, 2006, **42**, (7), pp. 1133–1145

10 Bhaumik, S., Srinivasan, M., Sadhu, S., Ghoshal, T.K.: 'Adaptive grid solution of risk-sensitive estimator problems'. Proc. IEEE INDICON, December 2005, pp. 356–360

11 Bhaumik, S., Srinivasan, M., Sadhu, S., Ghoshal, T.K.: 'A risk-sensitive estimator for nonlinear problems using the adaptive grid method'. IEEE Nonlinear Statistical Signal Processing Workshop (NSSPW), September 2006, pp. 1–4

12 Ristic, B., Arulampalam, S., Gordon, N.: 'Beyond the Kalman filter' (Artech House, USA, 2004)

13 Ito, K., Xiong, K.: 'Gaussian filters for nonlinear filtering problems', *IEEE Trans. Autom. Control*, 2000, **45**, (5), pp. 910–927

14 Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.: 'A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking', *IEEE Trans. Signal Process.*, 2002, **50**, (2), pp. 174–188

15 Lin, X., Kirubarajan, T., Bar-Shalom, Y., Maskell, S.: 'Comparison of EKF, pseudomeasurement filter and particle filter for a bearing-only target tracking problem'. Proc. SPIE Conf. Signal and Data Processing of Small Targets, April 2002

16 Sadhu, S., Mondal, S., Srinivasan, M., Ghoshal, T.K.: 'Sigma point Kalman filter for bearing only tracking', *Signal Process.*, 2006, **86**, (12), pp. 3769–3777